

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model  
Run on: June 26, 2005, 17:28:33 ; Search time 104.5 Seconds  
(without alignments)  
853.733 Million cell updates/sec

Title: US-10-099-926-1660

Perfect score: 198

Sequence: 1 caggtaaaatgactgttc.....tggtctatcggtagggtgc 116

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1717557 seqs, 384547976 residues

Total number of hits satisfying chosen parameters: 3435114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_n2p.model -DRV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US10099926/runat\_26062005\_142229\_22992/app.query.fasta\_1.263  
-DB=Published Applications AA -QFMT=faстан -SUFFIX=rapb -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=ext -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=30 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FCGAPOP=6 -FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10F\_PUBCOMB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US10G\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	58	29.3	93	16	US-10-425-115-267620
2	56	27.2	271	16	US-10-437-963-154451
3	55.5	26.9	521	16	US-10-437-963-200483
4	55	26.7	204	9	US-09-811-284-195
5	55	26.7	662	15	US-10-282-122A-67020
6	54	26.2	97	16	US-10-437-963-152525
7	53.5	26.0	129	16	US-10-425-115-236359
8	53	26.8	478	14	US-10-238-075-657
9	53	25.7	9234	10	US-09-942-025-13
10	52.5	25.5	62	16	US-10-437-963-119798
11	52	25.2	280	15	US-10-282-132A-47141
12	52	25.2	509	9	US-09-903-068-4
13	52	25.2	509	9	US-09-874-628-6
14	52	25.2	509	9	US-09-982-543A-4
15	52	25.2	509	14	US-10-005-228-2
16	52	25.2	509	15	US-10-170-385-361
17	52	25.2	509	16	US-10-600-645-6
18	52	25.2	509	17	US-10-739-413-4
19	52	25.2	561	16	US-10-437-963-178924
20	52	25.2	592	16	US-10-128-558-158
21	52	25.2	638	14	US-10-156-761-12352
22	52	25.2	820	16	US-10-437-963-199804
23	51.5	25.0	82	16	US-10-425-115-335400
24	51.5	25.0	357	9	US-09-947-027-9
25	51.5	25.0	357	13	US-10-091-009-9
26	51.5	25.0	357	15	US-10-424-599-220611
27	51.5	25.0	358	14	US-10-357-886-30
28	51.5	25.0	358	15	US-10-424-599-220612
29	51.5	25.0	359	15	US-10-425-114-52673
30	51.5	25.0	369	15	US-10-425-114-45888
31	51.5	25.0	390	15	US-10-425-114-51707
32	51.5	25.0	576	15	US-10-369-493-22649
33	51	24.8	36	14	US-10-133-128-128
34	51	24.8	36	14	US-10-289-660-128
35	51	24.8	36	16	US-10-693-057-128
36	51	24.8	36	17	US-10-693-056-128
37	51	24.8	36	17	US-10-840-723-128
38	51	24.8	36	17	US-10-871-602-128
39	51	24.8	42	9	US-09-864-761-38734
40	51	24.8	229	16	US-10-425-115-319896
41	51	24.8	790	16	US-10-437-963-170342
42	51	24.8	799	16	US-10-425-115-205000
43	51	24.8	4655	16	US-10-741-601-314
44	51	24.8	4655	16	US-10-479-875-3
45	51	24.8	4655	17	US-10-741-600-897

#### ALIGNMENTS

#### RESULT 1

US-10-425-115-267620  
: Sequence 267620, Application US/10425115  
: Publication No. US20040214272A1  
: GENERAL INFORMATION:  
: APPLICANT: La Rosa, Thomas J.  
: APPLICANT: Kovalic, David K.  
: APPLICANT: Zhou, Yihua  
: APPLICANT: Cao, Yongwei  
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
: TITLE OF INVENTION: Plance  
: FILE REFERENCE: 38-21(53222)B  
: CURRENT APPLICATION NUMBER: US/10/425,115  
: CURRENT FILING DATE: 2003-04-28  
: NUMBER OF SEQ ID NOS: 369326  
: SEQ ID NO 267620  
: LENGTH: 93  
: TYPE: PRT  
: ORGANISM: Zea mays  
: FEATURE:

### RESULT 3

Result No.	Score	Query		Length	DB	ID	Description
		Match					
C 1	56	27.2	165	4	US-09-250-991A-30108		Sequence 30108, A
C 2	55	27.7	196	4	US-09-270-767-61352		Sequence 61352, A
C 3	55	28.7	307	4	US-09-270-767-45825		Sequence 45825, A
C 4	52	25.2	303	3	US-08-158-735A-14		Sequence 14, Appl
C 5	52	25.2	465	4	US-09-949-016-9259		Sequence 9259, Ap
C 6	52	25.2	509	1	US-08-149-105-14		Sequence 14, Appl
C 7	52	25.2	509	1	US-08-317-847-14		Sequence 14, Appl
C 8	52	25.2	509	2	US-08-481-337A-4		Sequence 4, Appl
C 9	52	25.2	509	3	US-09-382-215-4		Sequence 4, Appl
C 10	52	25.2	509	3	US-09-395-116-4		Sequence 4, Appl
C 11	52	25.2	509	3	US-08-123-934A-6		Sequence 6, Appl
C 12	52	25.2	509	3	US-08-436-265-4		Sequence 4, Appl



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: June 26, 2005, 17:07:42 ; Search time 113.5 Seconds  
(without alignments)  
790.559 Million cell updates/sec

Title: US-10-099-926-1660  
Perfect score: 198  
Sequence: 1 caggtataaatgactgttc.....tggtctatcggtagggtgc 116

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0  
Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US10099926/runat\_26062005\_142225\_22915/app\_query.fasta\_1.263  
-DB=A\_Geneseq -QMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-LOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10099926@cgn\_1\_1\_224@runat\_26062005\_142225\_22915 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DLEXT=7

Database : A\_Geneseq\_16Dec04.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	61	29.6	246	5	ABP26906 Streptococcus agalactiae
C 2	59	28.6	71	4	AU39720 Propionib
C 3	59	28.6	71	6	ABM36239 Propionib
C 4	57	27.7	380	6	ABM67809 Photorhab
C 5	56	27.2	165	7	ABO81362 Pseudomon
C 6	55	26.7	204	4	AU29446 Human G p
C 7	55	26.7	204	5	ABG60734 Novel G p
C 8	55	26.7	638	4	ABG57754 Drosophil
C 9	55	26.7	662	6	ABU39096 Protein e
C 10	53	26.8	478	4	ABBS2636 Escherich

C 11	52	25.2	280	6	ABU19217	Abu19217 protein e
C 12	52	25.2	369	5	ABB93442	Abb93442 Herbicida
C 13	52	25.2	439	2	AAW79756	AAW79756 Euphorbia
C 14	52	25.2	509	2	AAR41920	AAR41920 MIR1.3/
C 15	52	25.2	509	2	AAR55367	AAR55367 Human Act
C 16	52	25.2	509	2	AAR85206	AAR85206 Human ALK
C 17	52	25.2	509	2	AAR70239	AAR70239 Serine/th
C 18	52	25.2	509	2	AAW03759	AAW03759 Mullerian
C 19	52	25.2	509	2	AAR94601	AAR94601 TAR-1 pol
C 20	52	25.2	509	2	AAV33300	AAV33300 Human hAL
C 21	52	25.2	509	5	ABG70732	ABG70732 Type I hu
C 22	52	25.2	509	5	ABP65198	ABP65198 Hypoxia-r
C 23	52	25.2	509	7	ADP57123	ADP57123 Rat Prote
C 24	52	25.2	509	7	ADE57125	ADE57125 Human Pro
C 25	52	25.2	509	8	ADQ89880	ADQ89880 Antagonis
C 26	52	25.2	592	7	ADE08003	ADE08003 Novel pro
C 27	51.5	26.0	126	4	ABB67055	ABB67055 Drosophil
C 28	51.5	25.0	151	4	AAU62123	AAU62123 Propionib
C 29	51.5	25.0	151	6	ABM58642	ABM58642 Propionib
C 30	51.5	25.0	181	3	AAAB36449	AAAB36449 Strawberry
C 31	51.5	25.0	181	3	AAV79666	AAV79666 Strawberry
C 32	51.5	25.0	357	5	AAU80016	AAU80016 Coniferyl
C 33	51.5	25.0	357	7	ADD93903	ADD93903 Quaking a
C 34	51.5	25.0	358	6	ADA38405	ADA38405 Soy bean
C 35	51.5	25.0	358	7	ADG88755	ADG88755 Soy bean
C 36	51.5	25.0	576	8	ADS44219	ADS44219 Bacterial
C 37	51.5	25.0	1764	6	ABR53481	ABR53481 Protein s
C 38	51.5	25.0	1764	7	ADK64566	ADK64566 Disease t
C 39	51	24.8	36	6	ABU61330	ABU61330 Human A d
C 40	51	24.8	36	6	ABU61409	ABU61409 Low densi
C 41	51	24.8	36	8	ADP21552	ADP21552 Human LDL
C 42	51	24.8	42	4	AAU19054	AAU19054 Peptide #
C 43	51	24.8	42	4	ABB38259	ABB38259 Peptide #
C 44	51	24.8	42	4	AAU31693	AAU31693 Peptide #
C 45	51	24.8	42	4	ABB23436	ABB23436 Protein #

ALIGNMENTS

RESULT 1  
ABP26906  
ID ABP26906 standard; protein; 246 AA.

XX AC

XX ABP26906;

XX DT 02-JUL-2002 (first entry)

XX DE Streptococcus polypeptide SEQ ID NO 2988.

XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

XX KW group A streptococcus; Streptococcus pyogenes; antibacterial;

XX KW antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX OS Streptococcus agalactiae.

XX PN WO200234771-A2.

XX XX 02-MAY-2002.

XX PF 29-OCT-2001; 2001WO-GB004789.

XX PR 27-OCT-2000; 2000GB-00026333.

XX PR 24-NOV-2000; 2000GB-00028727.

XX PR 07-MAR-2001; 2001GB-00005640.

XX (CHIR-) CHIRON SPA.

XX (GENO-) INST GENOMIC RES.

XX Telford J, Maignani V, Margarit Y Rosi, Grandi G, Fraser C;

XX Tettelin H;

XX WPI; 2002-352536/38.

XX DR N-PSDB; ABN67537.

XX New Streptococcus protein for the treatment or prevention of infection or  
PT disease caused by Streptococcus bacteria, such as meningitis, and for  
PT detecting a compound that binds to the protein.

XX Claim 1; Page 3451; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B  
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in  
CC the specification. The proteins have antibacterial and antiinflammatory  
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
CC antibodies that bind (I) are used in the manufacture of medicaments for  
CC the treatment or prevention of infection or disease caused by  
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
CC biological sample. (I) is used to determine whether a compound binds to  
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
CC used as a vaccine or diagnostic composition. The disease caused by  
CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
CC acid encoding (I) may be used to recombinantly produce (I) and may be  
CC used in gene therapy. Antibodies to (I) are used for affinity  
CC chromatography, immunoassays, and distinguishing/identifying  
CC Streptococcus proteins

XX Sequence 246 AA;

Alignment Scores:  
Pred. No.: 5.96 Length: 246  
Score: 61.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 5  
Best Local Similarity: 66.67% Mismatches: 0  
Query Match: 29.61% Indels: 0  
DB: 5 Gaps: 0

US-10-099-926-1660 (1-116) x ABP26906 (1-246)

QY 113 ACCTACCGATAGACAGTGGCTCAGATCTTACTTGTCTCTGCTT 69  
DB 129 SerMetProIleGluGlnTrpMetGlnSerPheLeuLeuLeuLeu 143

RESULT 2

AAU39720  
ID AAU39720 standard; protein; 71 AA.

XX AAU39720;

XX 13-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #616.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypervitosis; osteomyelitis;  
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US012865.

XX 21-APR-2000; 2000US-0199047P.

XX 02-JUN-2000; 2000US-0208841P.

XX 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX

DR WPI: 2001-616774/71.

DR N-PSDB; AAS59508.

XX Propionibacterium acnes polypeptides and nucleic acids useful for  
PT vaccinating against and diagnosing infections, especially useful for  
PT treating acne vulgaris.

XX Example 1; SEQ ID NO 915; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 71 AA;

Alignment Scores:  
Pred. No.: 10.7 Length: 71  
Score: 59.00 Matches: 15  
Percent Similarity: 51.28% Conservative: 5  
Best Local Similarity: 38.46% Mismatches: 17  
Query Match: 28.64% Indels: 2  
DB: 4 Gaps: 1

US-10-099-926-1660 (1-116) x AAU39720 (1-71)

QY 112 CCCTACCGATAGACAGTGGCTCAGATCTTACTTGTCTCTGCTGCTCTT 59  
DB 27 ProSerArgArgSerAlaAlaSerAlaLeuProCysSerCysTrpAlaThrArgSerPro 46

QY 58 CCCAATCACTGGTCACTCTGACCTCTGAACTCTGAACTCTGAACTCTGAACTCTT 2  
DB 47 AlaThrArgTrpSerProThrProThrSerTrpLeuValAlaThrGlnTrpArgPro 65

RESULT 3

ABM36239

ID ABM36239 standard; protein; 71 AA.

XX ABM36239;

XX 20-OCT-2003 (first entry)

XX Propionibacterium acnes predicted ORF-encoded polypeptide #915.

XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;  
KW immunostimulant; immune response; vaccine.

XX Propionibacterium acnes.

XX WO2003033515-A1.

XX 24-APR-2003.

XX 11-OCT-2002; 2002WO-US032727.

XX 15-OCT-2001; 2001US-00978825.

XX

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: June 26, 2005, 17:08:32 ; Search time 113.5 Seconds  
(without alignments)  
1046.717 Million cell updates/sec

Title: US-10-099-926-1660

Perfect score: 198

Sequence: 1 caggtataaaatgactgttc.....tggtctatcgtaggtgtgc 116

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_n2p.model -DEV=xlip

-Q=/cgn2\_1/USPRO.spool\_p/US10099926/runat\_26062005\_142226\_22921/app\_query.fasta\_1.263

-DB=UniProt -QFMT=faatan -SUFFIX=rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0

-UNITs=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45

-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=pfco -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US10099926\_@CGN\_1\_1\_244\_@runat\_26062005\_142226\_22921 -NCPU=6 -ICPU=3

-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	63	30.6	1603	2 Q6YXP5	Q6YXP5 Physcomitrella
C 2	62.5	30.3	3123	2 Q7R2N1	Q7R2N1 giardia lam
C 3	61	29.6	246	2 Q8E1P4	Q8E1P4 streptococ
C 4	61	29.6	246	2 Q8E760	Q8E760 streptococ
C 5	60	29.1	1031	1 YCFO ANTFO	Q85b66 anthoceros
C 6	58	28.2	469	1 REC3 METJA	Q58387 methanococ
C 7	58	28.2	488	2 Q6NTX4	Q6ntx4 xenopus lae
C 8	57	27.7	189	2 Q7QF49	Q7qf49 anopheles g
C 9	57	27.7	231	2 Q9GR92	Q9gr92 artemia san
C 10	57	27.7	328	2 Q8V3G2	Q8v3g2 swinepox vi
C 11	57	27.7	374	2 Q7N6M7	Q7n6m7 photorhabdu
C 12	57	27.7	502	2 Q7R220	Q7rf20 plasmodium
C 13	56.5	27.4	369	2 Q7Q344	Q7q344 geobacter s
C 14	56.5	27.4	489	2 Q8AAA3	Q8aaa3 bacteroides
C 15	56	27.2	505	2 Q6P3G3	Q6p3g3 brachydanio
C 16	55.5	26.9	521	2 Q67VA9	Q67va9 oryza sativ

#### ALIGNMENTS

##### RESULT 1

Q6YXP5	143	2	Q8C7P9
ID	55	26.7	Q8L854
AC	55	26.7	Q8L854
Q6YXP5	55	26.7	Q8L854
PRELIMINARY;	55	26.7	Q8L854
PRT; 1603 AA.	55	26.7	Q8L854
DT 05-JUL-2004 (T-EMBLrel. 27, Created)	55	26.7	Q8L854
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)	55	26.7	Q8L854
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)	55	26.7	Q8L854
DE Hypochemical protein ycf1.	55	26.7	Q8L854
GN Names=ycf1;	55	26.7	Q8L854
OS Physcomitrella patens subsp. patens.	55	26.7	Q8L854
OG Chloroplast.	55	26.7	Q8L854
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;	55	26.7	Q8L854
OC Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.	55	26.7	Q8L854
OX NCBI_TaxID=145481;	55	26.7	Q8L854
RP [1]	55	26.7	Q8L854
RP SEQUENCE FROM N.A.	55	26.7	Q8L854
RA Sugita C., Kobayashi Y., Setsuyuki A., Sugita C., Sugita M.;	55	26.7	Q8L854
RT "Complete chloroplast DNA sequence of the moss Physcomitrella patens.;"	55	26.7	Q8L854
RT evidence for the loss and relocation of rpoA from the chloroplast to	55	26.7	Q8L854
RT the nucleus.;"	55	26.7	Q8L854
RL Nucleic Acids Res. 31:5324-5531 (2003).	55	26.7	Q8L854
RN [2]	55	26.7	Q8L854
RN SEQUENCE FROM N.A.	55	26.7	Q8L854
RX MEDLINE=22079025; PubMed=12084583; DOI=10.1016/S0167-4781(02)00346-9;	55	26.7	Q8L854
RA Miyata Y., Sugita C., Kobayashi Y., Hagiwara M., Sugita M.;	55	26.7	Q8L854
RT "Chloroplast ribosomal S14 protein transcript is edited to create a	55	26.7	Q8L854
RT translation initiation codon in the moss Physcomitrella patens.;"	55	26.7	Q8L854
RN [3]	55	26.7	Q8L854
RN Biochim. Biophys. Acta 1576:346-349 (2002).	55	26.7	Q8L854
RN SEQUENCE FROM N.A.	55	26.7	Q8L854
RA Miyata Y., Sugita M.;	55	26.7	Q8L854
RT "Tissue- and stage-specific RNA editing of rps14 transcripts in moss	55	26.7	Q8L854
RT (Physcomitrella patens) chloroplasts.;"	55	26.7	Q8L854
RN [4]	55	26.7	Q8L854
RN J. Plant Physiol. 0:0-0 (2004).	55	26.7	Q8L854
RN SEQUENCE FROM N.A.	55	26.7	Q8L854
RP Sugita C., Tsuruya K., Aoki S., Sugita M.;	55	26.7	Q8L854
RT "Chloroplast transformation in tobacco and the moss Physcomitrella	55	26.7	Q8L854
RT patens.;"	55	26.7	Q8L854
RT (in) Unknown A. (eds.);	55	26.7	Q8L854

Q8c7p9 mus musculus  
Q8l854 arabidopsis  
Q8v116 drosophila  
Q8c933 pasteuria  
Q8n810 thodopseudo  
Q8gny1 bradyrhizob  
P12221 marchantia  
Q6g3sl bartonella  
Q85l09 macaca fasc  
Q83135 mus musculus  
Q8b1f7 mus musculus  
Q822h3 mus musculus  
Q810d6 mus musculus  
Q7z102 paramacium  
Q8e8l3 shewanella  
Q8irf4 drosophila  
Q8wt8 homo sapien  
Q86jw7 dictyosteli  
Q8nkl1 corynebacte  
Q69lp9 oryza sativ  
Q6lfs0 plasmodium  
Q57320 xenopus lae  
Q42475 xenopus lae  
Q7x2b9 escherichia  
Q8ff63 escherichia  
Q8g2l9 brucella su  
Q8vfy9 brucella me  
Q71q15 trimeresu  
Q84r70 oryza sativ





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 26, 2005, 15:48:53 ; Search time 2499 Seconds  
(without alignments)

1766.889 Million cell updates/sec

Title: US-10-099-926-1660

Perfect score: 116

Sequence: 1 caggtataaatgactgttc.....tgctctatcgtaggggtgc 116

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.\*

1: gb\_est1.\*

2: gb\_est2.\*

3: gb\_est3.\*

4: gb\_est4.\*

5: gb\_est5.\*

6: gb\_est6.\*

7: gb\_est7.\*

8: gb\_est8.\*

9: gb\_est9.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	109.4	94.3	289	AA861486	AA861486 ak21d08.s
2	109.4	94.3	331	BP431786	BP431786 nab50d09.s
3	109.4	94.3	335	AI804038	AI804038 tc60e09.x
4	109.4	94.3	336	AA989252	AA989252 or99e11.s
5	109.4	94.3	341	BG055806	BG055806 naf84b12.
6	109.4	94.3	352	BM713530	BM713530 UI-E-BJO-
7	109.4	94.3	353	AA846216	AA846216 al83c01.s
8	109.4	94.3	357	AA923290	AA923290 om51b12.s
9	109.4	94.3	363	AA885240	AA885240 am34f10.s
10	109.4	94.3	364	AI126331	AI126331 qd82c10.x
11	109.4	94.3	369	AA439603	AA439603 xt12h03.x
12	109.4	94.3	372	AA922691	AA922691 oh90g05.s
13	109.4	94.3	372	AI077383	AI077383 oy87g11.x
14	109.4	94.3	373	AA279973	AA279973 z888d01.s
15	109.4	94.3	374	AI352245	AI352245 qrlf01.x
16	109.4	94.3	384	AI1262594	AI1262594 qk41g07.x
17	109.4	94.3	395	AI1122960	AI1122960 ok32d03.s
18	109.4	94.3	396	CN263981	CN263981 170006000
19	109.4	94.3	401	AA766356	AA766356 oa31e07.s
20	109.4	94.3	402	AI121455	AI121455 DKPzp762L
21	109.4	94.3	403	AA749131	AA749131 ny03f02.s
22	109.4	94.3	405	AA568277	AA568277 nl86h11.s
23	109.4	94.3	408	AA769269	AA769269 n32e12.s
24	109.4	94.3	408	AI290444	AI290444 qo60g01.x

## ALIGNMENTS

RESULT 1  
AA861486  
LOCUS  
DEFINITION  
3', mRNA sequence.  
ACCESSION  
AA861486  
VERSION  
AA861486.1 GI:2953626  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 289)  
AUTHORS  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TUMOR  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Unpublished (1997)  
JOURNAL  
Contact: Robert Strausberg, Ph.D.  
COMMENT  
Email: cgapbs@mail.nih.gov  
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldo, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Cloned distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/Link at:  
www.bio.llnl.gov/bbrp/image/image.html  
Insert Length: 855 Std Error: 0.00  
Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 144.  
Location/Qualifiers  
1. .289  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1406607"  
/sex="male"  
/lab\_host="DH10B"  
/clone\_lib="Soares testis NHT"  
/notes="Vector: pr7T3D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
was prepared from mRNA obtained from Clontech  
Laboratories, Inc., and primed with a Not I - oligo(dT)  
primer [5',  
TGTACCATCTGAGTGGAGCGCGCCCAATTTTITTTTTTTT 3']  
Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT7T3 vector. Library

25 109.4 94.3 417 1 AI628725  
26 109.4 94.3 419 1 AA115910  
27 109.4 94.3 419 1 AA565410  
28 109.4 94.3 424 5 BU739559  
29 109.4 94.3 429 2 AW051372  
30 109.4 94.3 430 7 N72015  
31 109.4 94.3 432 2 AW104287  
32 109.4 94.3 436 1 AA418424  
33 109.4 94.3 444 1 AI346091  
34 109.4 94.3 444 1 AA488435  
35 109.4 94.3 449 7 W04573  
36 109.4 94.3 455 1 AA724674  
37 109.4 94.3 456 2 BF446908  
38 109.4 94.3 461 1 AI125792  
39 109.4 94.3 462 1 AI086020  
40 109.4 94.3 463 1 AI751291  
41 109.4 94.3 469 1 AA604657  
42 109.4 94.3 470 6 CD369490  
43 109.4 94.3 476 7 CR744650  
44 109.4 94.3 487 1 AI625216  
45 109.4 94.3 488 1 AA713612

AI628725 cy78b04.x  
AA115910 zm79f12.s  
AA565410 nk41902.s  
BU739559 UI-E-E30-  
AW051372 wy84f02.x  
N72015 yz96e06.s1  
AW104287 xd71h03.x  
AA418424 zv92h11.s  
AI346091 qp43d08.x  
AA488435 ab39f05.s  
W04573 zb92c04.s1  
AA724674 aj12h05.s  
BF446908 7065d04.x  
AI125792 qd95g02.x  
AI086020 oy70f06.x  
AI751291 cn08g11.y  
AA604657 no80d10.s  
CD369490 UI-H-F71-  
CR744650 CR744650  
AI625216 ts24b06.x  
AA713612 nv70h02.s

went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo. "

```

ORIGIN
Query Match          94.3%; Score 109.4; DB 1; Length 289;
Best Local Similarity 99.1%; Pred. No. 8.3e-26;
Matches 110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 AAAAAATGACTGTTTCAGGAGTGTTCAGTAGGTCAGATGACCAAGTATGGGAATACCTT 65
Db 169 AAAAAATGACTGTTTCAGGAGTGTTCAGTAGGTCAGATGACCAAGTATGGGAATACCTT 228

QY 66 CGTAAGCAGGAGCAAGTAAGATCTGAGCCACTGTTCTATCGGTAGGGTGTC 116
Db 229 CGTAAGCAGGAGCAAGTAAGATCTGAGCCACTGTTCTATCGGTAGGGTGTC 279

RESULT 2
BF4311786
LOCUS
DEFINITION
nab50d09_x1 Soares NSF F8_9W OT_PA_P_S1 Homo sapiens cDNA clone
IMAGE:3269416 3', mRNA sequence.
ACCESSION
BF4311786
VERSION
BF4311786
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 331)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLM; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 298.
Location/Qualifiers
1..331
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3269416"
/lab_host="DH10B"
/clone_lib="Soares NSF F8_9W OT_PA_P_S1"
/notes="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and clones: Soares NB2HP pool 1:
309384-310919, 323208-325895 Soares NB2HP pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares NB2HF8-9W pool 1:
758280-760583, 772104-774407 Soares NBHPA pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NBHOT
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Bonaldo."

ORIGIN
Query Match          94.3%; Score 109.4; DB 2; Length 331;
Best Local Similarity 99.1%; Pred. No. 8.6e-26;
Matches 110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 AAAAAATGACTGTTTCAGGAGTGTTCAGTAGGTCAGATGACCAAGTATGGGAATACCTT 65
Db 165 AAAAAATGACTGTTTCAGGAGTGTTCAGTAGGTCAGATGACCAAGTATGGGAATACCTT 224

QY 66 CGTAAGCAGGAGCAAGTAAGATCTGAGCCACTGTTCTATCGGTAGGGTGTC 116
Db 225 CGTAAGCAGGAGCAAGTAAGATCTGAGCCACTGTTCTATCGGTAGGGTGTC 275

RESULT 4
AA989252
LOCUS
DEFINITION
or99e11.s1 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1604012 3',
mRNA sequence.
ACCESSION
AA989252
VERSION
AA989252.1 GI:3173874

```

```

QY 66 CGTAAGCAGGAGCAAGTAAGATCTGAGCCACTGTTCTATCGGTAGGGTGTC 116
Db 225 CGTAAGCAGGAGCAAGTAAGATCTGAGCCACTGTTCTATCGGTAGGGTGTC 275

RESULT 3
AI804038
LOCUS
DEFINITION
tc6e09_x1 Soares NHMPu_S1 Homo sapiens cDNA clone IMAGE:2069032
3', mRNA sequence.
ACCESSION
AI804038
VERSION
AI804038
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 335)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLM; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1213 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 332.
Location/Qualifiers
1..335
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2069032"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/clone_lib="Soares NHMPu S1"
/notes="Organ: mixed (see below); Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBHM, pregnant uterus
NBHPu, and fetal heart NBH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."

ORIGIN
Query Match          94.3%; Score 109.4; DB 1; Length 335;
Best Local Similarity 99.1%; Pred. No. 8.6e-26;
Matches 110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 AAAAAATGACTGTTTCAGGAGTGTTCAGTAGGTCAGATGACCAAGTATGGGAATACCTT 65
Db 165 AAAAAATGACTGTTTCAGGAGTGTTCAGTAGGTCAGATGACCAAGTATGGGAATACCTT 224

QY 66 CGTAAGCAGGAGCAAGTAAGATCTGAGCCACTGTTCTATCGGTAGGGTGTC 116
Db 225 CGTAAGCAGGAGCAAGTAAGATCTGAGCCACTGTTCTATCGGTAGGGTGTC 275

RESULT 4
AA989252
LOCUS
DEFINITION
or99e11.s1 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1604012 3',
mRNA sequence.
ACCESSION
AA989252
VERSION
AA989252.1 GI:3173874

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: June 26, 2005, 17:16:47 ; Search time 25 Seconds  
(without alignments)

892.892 Million cell updates/sec

Title: US-10-099-926-1660

Perfect score: 198

Sequence: 1 caggtaaaatgactgttc.....tgctctatcggtagggtgc 116

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame\_n2p.model -DSV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US10099926/runat 26062005 142227 22934/app\_query.fasta\_1.263  
-DB=PIR -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10099926@cgn 1 1 63 @runat 26062005 142227 22934 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG  
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR.79.\*

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	58	28.2	469	A64422	conserved hypothetical
C 2	55	26.7	1068	S01519	hypothetical prote
C 3	52.5	25.5	185	A13454	hypothetical prote
C 4	52	25.2	96	T22379	hypothetical prote
C 5	52	25.2	280	D70156	conserved hypothet
C 6	52	25.2	411	G89074	protein K04A8.5 [i
C 7	52	25.2	421	S33998	K421R protein - Af
C 8	52	25.2	509	A45992	activin A receptor
C 9	52	25.2	509	I59576	transforming growt
C 10	52	25.2	509	A49664	activin type I rec
C 11	51.5	25.0	201	F72454	hypothetical prote
C 12	51.5	25.0	357	S23525	cinnamyl-alcohol d
C 13	51.5	25.0	358	S31572	cinnamyl-alcohol d
C 14	51.5	25.0	432	AE2403	ATP-binding protei

C 15	51.5	25.0	551	2	A60047	adenyl cyclase-a
C 16	51.5	25.0	1764	2	S37827	hypothetical prote
C 17	51	24.8	284	2	B95306	probable regulator
C 18	51	24.8	455	2	E96817	hypothetical prote
C 19	50.5	25.5	72	2	H82519	hypothetical prote
C 20	50.5	24.5	357	2	S23526	cinnamyl-alcohol d
C 21	50.5	25.5	482	2	F96778	hydroxymethylpyrim
C 22	50	24.3	198	1	H64771	hypothetical prote
C 23	50	24.3	198	2	F90688	hypothetical prote
C 24	50	24.3	198	2	B85539	hypothetical prote
C 25	50	24.3	314	2	T31866	hypothetical prote
C 26	50	24.3	357	2	T05413	cinnamyl-alcohol d
C 27	50	24.3	387	2	F71279	probable sugar ABC
C 28	50	24.3	406	2	A83795	hypothetical prote
C 29	50	24.3	619	2	T64087	translation elonga
C 30	50	24.3	621	2	B95897	probable cellulose
C 31	50	24.3	1172	2	C70619	probable lyx prot
C 32	50	24.3	1443	2	T02491	probable ABC trans
C 33	50	24.3	1738	2	S20614	conserved hypothet
C 34	50	24.3	2091	2	A97077	hypothetical prote
C 35	50	25.3	2541	2	T29340	hypothetical prote
C 36	49.5	24.0	199	2	H70709	hypothetical prote
C 37	49.5	24.0	303	2	E69584	3-methyladenine DN
C 38	49.5	24.0	612	2	C90419	glucan 1,4 alpha g
C 39	49	23.8	197	2	F75436	conserved hypothet
C 40	49	23.8	296	2	A64110	cell division inhi
C 41	49	23.8	366	2	F86794	hypothetical prote
C 42	49	23.8	414	2	AG2195	hypothetical prote
C 43	49	24.7	810	2	S44653	F42H10.5 protein -
C 44	49	24.7	1072	2	AD1280	SNF2-type helicase
C 45	49	23.8	1197	2	D86317	protein F15H18.21

#### ALIGNMENTS

##### RESULT 1

A64422  
conserved hypothetical protein MJ0977 - Methanococcus jannaschii  
C/Species: Methanococcus jannaschii  
C/Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004  
C/Accession: A64422  
R/Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A/Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C  
A/Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A/Reference number: A64300; MUID:96337999; PMID:8688087  
A/Accession: A64422  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-469 <BUL>  
A/Cross-references: UNIPROT:Q58387; GB:U67540; GB:L77117; NID:gl591631; PIDN:AAB98980.1;  
C/Genetics:  
A/Map position: FOR905956-911005  
C/Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0977

Alignment Scores:				
Pred. No.:	3.81	Length:	469	
Score:	58.00	Matches:	8	
Percent Similarity:	70.59%	Conservative:	4	
Best Local Similarity:	47.06%	Mismatches:	5	
Query Match:	28.16%	Indels:	0	
DB:	2	Gaps:	0	
US-10-099-926-1660 (1-116) x A64422 (1-469)				
QY 94	GCCTCAGATCTTACTTGCCTTCCTTACGAAGTATTCCTCAATCACTGGTCA 44			
Db 261	GLYASGLULULULYSYSLEUVALPROASHISTPTPHR 277			
RESULT 2				
S01519				



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 26, 2005, 19:21:13 ; Search time 388 Seconds  
(without alignments)  
1868.711 Million cell updates/sec

Title: US-10-099-926-1660

Perfect score: 116

Sequence: 1 caggtataaataactgttc.....tggtctatcggtagggtgc 116

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6067389 seqs, 3125258755 residues

Total number of hits satisfying chosen parameters: 12134778

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*
- 20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq.\*
- 21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq.\*
- 22: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*
- 24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
- 25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	116	100.0	116	9	US-09-920-300A-1660
2	116	100.0	116	13	Sequence 1660, Ap
3	116	100.0	116	16	Sequence 1660, Ap
4	109.4	94.3	297	9	US-09-922-217-419
5	109.4	94.3	297	9	US-09-922-217-419
6	109.4	94.3	297	13	US-10-025-380-146
7	109.4	94.3	511	9	US-09-922-217-146

RESULT 1

US-09-920-300A-1660  
; Sequence 1660, Application US/09920300A

; Patent No. US20020136728A1

; GENERAL INFORMATION:

; APPLICANT: King, Gordon E.

; APPLICANT: Meagher, Madeleine Joy

; APPLICANT: Xu, Jiangchun

; APPLICANT: Secrist, Heather

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.547

; CURRENT APPLICATION NUMBER: US/09/920,300A

; NUMBER OF SEQ ID NOS: 1789

; SOFTWARE: FASTSEQ for Windows Version 4.0

; SEQ ID NO 1660

; LENGTH: 116

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-920-300A-1660

Query Match 100.0%; Score 116; DB 9; Length 116;

Best Local Similarity 100.0%; Pred. No. 6.6e-32;

Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGTAAAAATGACTGTTTCAGAGTGTTCAGTAGGTTCAAGTAGGTCAGATGACCATGATTGGAA 60

DB 1 CAGGTAAAAATGACTGTTTCAGAGTGTTCAGTAGGTTTCAGATGACCATGATTGGAA 60

Sequence 146, App  
Sequence 146, App  
Sequence 5876, Ap  
Sequence 1606, Ap  
Sequence 641, App  
Sequence 374, App  
Sequence 1982, Ap  
Sequence 4302, A  
Sequence 4302, A  
Sequence 55690, A  
Sequence 255081, A  
Sequence 255081, A  
Sequence 60470, A  
Sequence 60471, A  
Sequence 60471, A  
Sequence 60471, A  
Sequence 123634, A  
Sequence 59314, A  
Sequence 722, App  
Sequence 612, App  
Sequence 612, App  
Sequence 13, Appl  
Sequence 22310, A  
Sequence 131358, A  
Sequence 62, Appl  
Sequence 1496, Ap  
Sequence 1627, Ap  
Sequence 14724, A  
Sequence 171, App  
Sequence 171, App  
Sequence 45379, A  
Sequence 56947, A  
Sequence 180, App  
Sequence 292399, A  
Sequence 292399, A  
Sequence 119053, A  
Sequence 17756, A  
Sequence 157933, A

ALIGNMENTS

QY 61 TACTTCGTAAGCAGGAGCAAGTAAGATCTCGAGCCACTGTTCTATCGGTAGGGTGTC 116  
|||||  
Db 61 TACTTCGTAAGCAGGAGCAAGTAAGATCTCGAGCCACTGTTCTATCGGTAGGGTGTC 116  
|||||

## RESULT 2

US-10-033-528-1660  
; Sequence 1660, Application US/10033528  
; Publication No. US20020131971A1  
; GENERAL INFORMATION:  
; APPLICANT: King, Gordon E.  
; APPLICANT: Meagher, Madeleine Joy  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Secrist, Heather  
; APPLICANT: Wang, Tongtong  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.547C1  
; CURRENT APPLICATION NUMBER: US/10/033,528  
; CURRENT FILING DATE: 2001-12-26  
; NUMBER OF SEQ ID NOS: 1896  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1660  
; LENGTH: 116  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-033-528-1660

Query Match 100.0%; Score 116; DB 13; Length 116;  
Best Local Similarity 100.0%; Pred. No. 6.6e-32;  
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGTAAAAATGACTGTTCCAGGAGTGTTCAGGTAGGTCAGATGACCACTGATTGGAA 60  
|||||  
Db 1 CAGGTAAAAATGACTGTTCCAGGAGTGTTCAGGTAGGTCAGATGACCACTGATTGGAA 60  
|||||  
QY 61 TACTTCGTAAGCAGGAGCAAGTAAGATCTCGAGCCACTGTTCTATCGGTAGGGTGTC 116  
|||||  
Db 61 TACTTCGTAAGCAGGAGCAAGTAAGATCTCGAGCCACTGTTCTATCGGTAGGGTGTC 116  
|||||

## RESULT 3

US-10-099-926-1660  
; Sequence 1660, Application US/10099926  
; Publication No. US20030166064A1  
; GENERAL INFORMATION:  
; APPLICANT: King, Gordon E.  
; APPLICANT: Meagher, Madeleine Joy  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Secrist, Heather  
; APPLICANT: Jiang, Yuqiu  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.547C2  
; CURRENT APPLICATION NUMBER: US/10/099,926  
; CURRENT FILING DATE: 2002-03-17  
; NUMBER OF SEQ ID NOS: 1982  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1660  
; LENGTH: 116  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-099-926-1660

Query Match 100.0%; Score 116; DB 16; Length 116;  
Best Local Similarity 100.0%; Pred. No. 6.6e-32;  
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGTAAAAATGACTGTTCCAGGAGTGTTCAGGTAGGTCAGATGACCACTGATTGGAA 60  
|||||  
Db 1 CAGGTAAAAATGACTGTTCCAGGAGTGTTCAGGTAGGTCAGATGACCACTGATTGGAA 60  
|||||  
QY 61 TACTTCGTAAGCAGGAGCAAGTAAGATCTCGAGCCACTGTTCTATCGGTAGGGTGTC 116  
|||||

Db 61 TACTTCGTAAGCAGGAGCAAGTAAGATCTCGAGCCACTGTTCTATCGGTAGGGTGTC 116  
|||||

## RESULT 4

US-09-922-217-419/c  
; Sequence 419, Application US/09922217  
; Patent No. US20020076414A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Meagher, Madeleine Joy  
; APPLICANT: Stolk, John A.  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Smith, Carole Lynn  
; APPLICANT: King, Gordon E.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Clapper, Jonathan D.  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS  
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.471C13  
; CURRENT APPLICATION NUMBER: US/09/922,217  
; CURRENT FILING DATE: 2001-08-03  
; NUMBER OF SEQ ID NOS: 1124  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 419  
; LENGTH: 297  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-922-217-419

Query Match 94.3%; Score 109.4; DB 9; Length 297;  
Best Local Similarity 99.1%; Pred. No. 2.3e-29;  
Matches 110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 AAAAAATGACTGTTCCAGGAGTGTTCAGGTAGGTCAGATGACCACTGATTGGAACTTT 65  
|||||  
Db 136 AAAAAATGACTGTTCCAGGAGTGTTCAGGTAGGTCAGATGACCACTGATTGGAACTTT 77  
|||||  
QY 66 CGTAAGCAGGAGCAAGTAAGATCTCGAGCCACTGTTCTATCGGTAGGGTGTC 116  
|||||  
Db 76 CGTAAGCAGGAGCAAGTAAGATCTCGAGCCACTGTTCTATCGGTAGGGTGTC 26  
|||||

## RESULT 5

US-09-833-263-419/c  
; Sequence 419, Application US/09833263  
; Patent No. US20020110547A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Aijun  
; APPLICANT: Clapper, Jonathan D.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Meagher, Madeleine J.  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.471C12  
; CURRENT APPLICATION NUMBER: US/09/833,263  
; CURRENT FILING DATE: 2001-04-10  
; NUMBER OF SEQ ID NOS: 1093  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 419  
; LENGTH: 297  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-833-263-419

Query Match 94.3%; Score 109.4; DB 9; Length 297;  
Best Local Similarity 99.1%; Pred. No. 2.3e-29;  
Matches 110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 AAAAAATGACTGTTCCAGGAGTGTTCAGGTAGGTCAGATGACCACTGATTGGAACTTT 65  
|||||

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 26, 2005, 13:48:37 ; Search time 1711 Seconds  
(without alignments)  
3285.099 Million cell updates/sec

Title: US-10-099-926-1660

Perfect score: 116

Sequence: 1 caggtataaaatgactgttc.....tggtctatcgtaggggtgc 116

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_btg.\*

3: gb\_in.\*

4: gb\_on.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_ey.\*

13: gb\_un.\*

14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	100.0	116	6	AX397445 Sequence
C 2	109.4	94.3	297	6	BD265413 Compounds
C 3	109.4	94.3	297	6	AX192852 Sequence
C 4	109.4	94.3	511	6	BD265147 Compounds
C 5	109.4	94.3	511	6	AX401133 Sequence
C 6	109.4	94.3	511	6	AX192579 Sequence
C 7	109.4	94.3	518	6	AX150126 Sequence
C 8	109.4	94.3	819	11	G06474 human STS W
C 9	109.4	94.3	1748	9	AF054988 Homo sapi
C 10	109.4	94.3	3594	9	AX408959 Sequence
C 11	109.4	94.3	3594	9	DI3636 Homo sapien
C 12	109.4	94.3	3600	9	BC020981 Homo sapi
C 13	109.4	94.3	118847	9	AC013413 Homo sapi
C 14	107.8	92.9	761	11	BV209266 GTF3C2_33
C 15	43.8	37.8	3560	10	BC034369 Mus muscu
C 16	43.8	37.8	4301	10	AK129033 Mus muscu
C 17	43.8	37.8	4861	10	BC043100 Mus muscu
C 18	43.8	37.8	204468	10	AC114619 Mus muscu
C 19	43.8	37.8	209211	2	AC084883 Mus muscu

C 20	43.8	37.8	216195	10	AC109608
C 21	34.8	30.0	255272	2	AC097691
C 22	33.2	28.6	188946	10	AC123747
C 23	33	28.4	127151	5	BX842593
C 24	33	28.4	178005	2	EX649315
C 25	33	28.4	180048	5	AL772163
C 26	33	28.4	196931	5	BX284619
C 27	33	28.4	216967	2	CR556708
C 28	33	28.4	221163	2	EX572628
C 29	33	28.4	240212	2	EX640547
C 30	33	28.4	252440	5	EX005425
C 31	33	28.4	272535	5	CR4505212
C 32	33	28.4	273751	2	CR450700
C 33	32.6	28.1	185879	2	AC109019
C 34	32.6	28.1	229820	2	AC121305
C 35	32	27.6	154065	2	AC084079
C 36	32	27.6	154539	2	AC079187
C 37	32	27.6	156909	9	AB020863
C 38	32	27.6	168871	2	AC016117
C 39	32	27.6	171075	2	AC026842
C 40	32	27.6	186218	9	AP006248
C 41	31.8	27.4	210858	2	AC129368
C 42	31.8	27.4	225767	2	AC098011
C 43	31.8	27.4	255368	2	AC109116
C 44	31.6	27.2	139695	5	EX324128
C 45	31.4	27.1	169861	2	CR354374

## ALIGNMENTS

RESULT 1  
AX397445  
LOCUS AX397445  
DEFINITION Sequence 1660 from Patent WO0212328.  
ACCESSION AX397445  
VERSION AX397445.1  
KEYWORDS GI:21068192  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1  
AUTHORS King, G.E., Meagher, M.J., Xu, J. and Secrist, H.  
TITLE Compositions and methods for the therapy and diagnosis of colon cancer  
JOURNAL Patent: WO 0212328-A.1660 14-FEB-2002;  
CORIXA CORPORATION (US)  
FEATURES  
Location/Qualifiers  
1..116  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

Query Match 100.0%; Score 116; DB 6; Length 116;  
Best Local Similarity 100.0%; Pred No. 1.7e-29;  
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGTAAAAATGACTGTTTCAGGAGTGTTCAGTAGGGTCAGATGACCACTGTTCTATCGTAGGGTGTTC 116  
1 CAGGTAAAAATGACTGTTTCAGGAGTGTTCAGTAGGGTCAGATGACCACTGTTCTATCGTAGGGTGTTC 116

Db 1 CAGGTAAAAATGACTGTTTCAGGAGTGTTCAGTAGGGTCAGATGACCACTGTTCTATCGTAGGGTGTTC 116  
1 CAGGTAAAAATGACTGTTTCAGGAGTGTTCAGTAGGGTCAGATGACCACTGTTCTATCGTAGGGTGTTC 116

QY 61 TACTTCGTAGCAGGAGCAGTAGTAAGTCTGAGCCACTGTTCTATCGTAGGGTGTTC 116  
1 TACTTCGTAGCAGGAGCAGTAGTAAGTCTGAGCCACTGTTCTATCGTAGGGTGTTC 116

Db 61 TACTTCGTAGCAGGAGCAGTAGTAAGTCTGAGCCACTGTTCTATCGTAGGGTGTTC 116  
1 TACTTCGTAGCAGGAGCAGTAGTAAGTCTGAGCCACTGTTCTATCGTAGGGTGTTC 116

RESULT 2  
BD265413/c  
LOCUS BD265413  
DEFINITION Compounds for immunotherapy and diagnosis of colonic cancer and method of using the same.

QY 1 CAGGTAAAAATGACTGTTTCAGGAGTGTTCAGTAGGGTCAGATGACCACTGTTCTATCGTAGGGTGTTC 116  
1 CAGGTAAAAATGACTGTTTCAGGAGTGTTCAGTAGGGTCAGATGACCACTGTTCTATCGTAGGGTGTTC 116

Db 1 CAGGTAAAAATGACTGTTTCAGGAGTGTTCAGTAGGGTCAGATGACCACTGTTCTATCGTAGGGTGTTC 116  
1 CAGGTAAAAATGACTGTTTCAGGAGTGTTCAGTAGGGTCAGATGACCACTGTTCTATCGTAGGGTGTTC 116

QY 61 TACTTCGTAGCAGGAGCAGTAGTAAGTCTGAGCCACTGTTCTATCGTAGGGTGTTC 116  
1 TACTTCGTAGCAGGAGCAGTAGTAAGTCTGAGCCACTGTTCTATCGTAGGGTGTTC 116

Db 61 TACTTCGTAGCAGGAGCAGTAGTAAGTCTGAGCCACTGTTCTATCGTAGGGTGTTC 116  
1 TACTTCGTAGCAGGAGCAGTAGTAAGTCTGAGCCACTGTTCTATCGTAGGGTGTTC 116

RESULT 2  
BD265413/c  
LOCUS BD265413  
DEFINITION Compounds for immunotherapy and diagnosis of colonic cancer and method of using the same.

ACCESSION BD265413  
VERSION BD265413.1 GI:33075181  
KEYWORDS JP 2002533082-A/411.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS Xu, J., Lodes, M.J., Secrist, H., Benson, D.R., Meagher, M.J., Stolk, J.,  
Wang, T. and Yuqiu, J.  
TITLE Compounds for immunotherapy and diagnosis of colonic cancer and  
method of using the same  
JOURNAL Patent: JP 2002533082-A 411 08-OCT-2002;  
CORIXA CORP  
COMMENT OS Homo sapiens (human)  
PN JP 2002533082-A/411  
PD 08-OCT-2002  
PF 23-DEC-1999 JP 2000589697  
PR 23-DEC-1998 US 09/221298, 02-JUL-1999 US 09/347496 PR  
22-SEP-1999 US 09/401064, 19-NOV-1999 US 09/444242 PR  
02-DEC-1999 US 09/454150  
PI JIANGCHUN XU, MICHAEL J LODES, HEATHER SECRIST, DARIN R BENSON,  
PI MADELEINE JOY MEAGHER, JOHN STOLK, TONGTONG WANG, JIANG YUQIU PC  
C12N15/09, A61K31/711, A61K35/14, A61K38/00, A61K39/00, A61K39/395, PC  
A61K39/395,  
PC A61P35/00, C07K14/47, C07K16/18, C07K19/00, C12N1/15, C12N1/19, PC  
C12N1/21,  
PC  
C12N5/06, C12N5/10, C12Q1/02, C12Q1/68, G01N33/53, G01N33/53, G01N33/ PC  
566,  
G01N33/574, G01N33/577, G01N33/58, C12N15/00, C12N5/00, C12N5/00,  
PC A61K37/02  
CC Compounds for immunotherapy and diagnosis of colonic cancer  
CC using the same  
CC Key Location/Qualifiers  
FH source 1..297  
FT /organism="Homo sapiens (human)".  
FT Location/Qualifiers  
1..297  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 94.3%; Score 109.4; DB 6; Length 297;  
Best Local Similarity 99.1%; Pred. No. 3.7e-27;  
Matches 110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 6 AAAAAATGACTGTTTCAGGAGTGTTCAGTAGGGTTCAGATGACCATGATGGGAATACCTT 65  
DB 136 AAAACATGACTGTTTCAGGAGTGTTCAGTAGGGTTCAGATGACCATGATGGGAATACCTT 77  
QY 66 CGTAAGCAGGAGCAAGTAAGATCTGAGCCACTGTTCTATCGGTAGGGTGTGC 116  
DB 76 CGTAAGCAGGAGCAAGTAAGATCTGAGCCACTGTTCTATCGGTAGGGTGTGC 26

RESULT 3  
AX192852/c  
LOCUS AX192852 419 bp DNA linear PAT 15-AUG-2001  
DEFINITION Sequence 419 from Patent WO0149716.  
ACCESSION AX192852  
VERSION AX192852.1 GI:15210808  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS Xu, J., Lodes, M.J., Secrist, H., Benson, D.R., Meagher, M.J.,  
Stolk, J.A., King, G.E., Wang, T. and Jiang, Y.  
TITLE Compounds for immunotherapy and diagnosis of colon cancer and

methods for their use  
Patent: WO 0149716-A 419 12-JUL-2001;  
CORIXA CORPORATION (US)  
FEATURES  
source Location/Qualifiers  
1..297  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 94.3%; Score 109.4; DB 6; Length 297;  
Best Local Similarity 99.1%; Pred. No. 3.7e-27;  
Matches 110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 6 AAAAAATGACTGTTTCAGGAGTGTTCAGTAGGGTTCAGATGACCATGATGGGAATACCTT 65  
DB 136 AAAACATGACTGTTTCAGGAGTGTTCAGTAGGGTTCAGATGACCATGATGGGAATACCTT 77  
QY 66 CGTAAGCAGGAGCAAGTAAGATCTGAGCCACTGTTCTATCGGTAGGGTGTGC 116  
DB 76 CGTAAGCAGGAGCAAGTAAGATCTGAGCCACTGTTCTATCGGTAGGGTGTGC 26

RESULT 4  
BD265147  
LOCUS BD265147  
DEFINITION Compounds for immunotherapy and diagnosis of colonic cancer and  
method of using the same.  
ACCESSION BD265147  
VERSION BD265147.1 GI:33074915  
KEYWORDS JP 2002533082-A/145.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS Xu, J., Lodes, M.J., Secrist, H., Benson, D.R., Meagher, M.J., Stolk, J.,  
Wang, T. and Yuqiu, J.  
TITLE Compounds for immunotherapy and diagnosis of colonic cancer and  
method of using the same  
JOURNAL Patent: JP 2002533082-A 145 08-OCT-2002;  
CORIXA CORP  
COMMENT OS Homo sapiens (human)  
PN JP 2002533082-A/145  
PD 08-OCT-2002  
PF 23-DEC-1999 JP 2000589697  
PR 23-DEC-1998 US 09/221298, 02-JUL-1999 US 09/347496 PR  
22-SEP-1999 US 09/401064, 19-NOV-1999 US 09/444242 PR  
02-DEC-1999 US 09/454150  
PI JIANGCHUN XU, MICHAEL J LODES, HEATHER SECRIST, DARIN R BENSON,  
PI MADELEINE JOY MEAGHER, JOHN STOLK, TONGTONG WANG, JIANG YUQIU PC  
C12N15/09, A61K31/711, A61K35/14, A61K38/00, A61K39/00, A61K39/395, PC  
A61K39/395,  
PC A61P35/00, C07K14/47, C07K16/18, C07K19/00, C12N1/15, C12N1/19, PC  
C12N1/21,  
PC  
C12N5/06, C12N5/10, C12Q1/02, C12Q1/68, G01N33/53, G01N33/53, G01N33/ PC  
566,  
G01N33/574, G01N33/577, G01N33/58, C12N15/00, C12N5/00, C12N5/00,  
PC A61K37/02  
CC Compounds for immunotherapy and diagnosis of colonic cancer  
CC using the same  
CC Key Location/Qualifiers  
FH source 1..511  
FT /organism="Homo sapiens (human)".  
FT Location/Qualifiers  
1..511  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 26, 2005, 13:26:17 ; Search time 286 Seconds  
(without alignments)

2401.014 Million cell updates/sec

Title: US-10-099-926-1660

Perfect score: 116

Sequence: 1 caggtaaaaaaactgttc.....tggtctatcggtagggtgc 116

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

1: Geneseqn1980s:\*

2: Geneseqn1990s:\*

3: Geneseqn2000s:\*

4: Geneseqn2001s:\*

5: Geneseqn2001bs:\*

6: Geneseqn2002as:\*

7: Geneseqn2002bs:\*

8: Geneseqn2003as:\*

9: Geneseqn2003bs:\*

10: Geneseqn2003cs:\*

11: Geneseqn2003ds:\*

12: Geneseqn2004as:\*

13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	100.0	116	6	ABK46109
C 2	109.4	94.3	297	3	AA78132
C 3	109.4	94.3	297	4	AA128870
C 4	109.4	94.3	297	8	ABZ33056
5	109.4	94.3	511	3	AA77866
6	109.4	94.3	511	4	AA128604
7	109.4	94.3	511	8	AAZ32790
8	109.4	94.3	518	4	AAH50772
C 9	109.4	94.3	3594	6	ABN95108
C 10	109.4	94.3	3594	13	ACN37658
C 11	108.4	93.4	395	7	ADS71777
C 12	93.4	80.5	410	13	ADQ78802
C 13	76.2	65.7	251	2	AA722211
C 14	29.2	25.2	618	5	ABV59295
C 15	29.2	25.2	1515	3	AA373324
16	29	25.0	2721	4	AA63796
17	29	25.0	3710	8	ABZ35917
18	29	25.0	5778	10	ABE31387
19	28.8	24.8	231	3	AA57239
20	28.8	24.8	231	6	ABT12562

21	28.8	24.8	231	10	ACD91856	Ac91856 Human col
22	28.8	24.8	1404	6	ABZ12807	Abz12807 Arabidops
23	28.8	24.8	1404	12	ADJ98175	Adj98175 Thale cfe
24	28.8	24.8	1479	3	AAZ88979	Aaz88979 A. thalia
25	28.6	24.7	425	8	ABX36331	Abx36331 Bovine ES
26	28.4	24.5	3951	12	ADQ86892	Adq86892 Human tum
27	28.4	24.5	111084	12	ADQ18808	Adq18808 Human sof
C 28	28.2	24.3	524	13	ACN59943	Acn59943 Cotton gy
C 29	28.2	24.3	909	2	AAV37170	Aav37170 DNA seque
C 30	28.2	24.3	909	4	AAH01761	Aah01761 Wolinella
31	28.2	24.3	3054	13	ADT46941	Adt46941 Bacterial
C 32	28.2	24.3	265118	5	AAH41227	Aah41227 Pyrococcu
33	28	24.1	353	5	ABA20723	Ab20723 Human ner
34	28	24.1	353	5	ABA20724	Ab20724 Human ner
C 35	27.8	24.0	315	2	AAV87702	Aav87702 EST clone
36	27.8	24.0	9042	5	ABA21227	Ab21227 Human ner
37	27	23.3	614	10	ABX12493	Abx12493 CDNA enco
38	27	23.3	209612	12	ADQ59395	Adq59395 Human can
C 39	26.8	23.1	742	10	ADB67959	Adb67959 Human lun
C 40	26.8	23.1	909	8	ACA42953	Aca42953 Prokaryot
C 41	26.8	23.1	1840	4	AAK94394	Aak94394 Human ful
C 42	26.8	23.1	1840	12	ADL31110	Adl31110 Full leng
43	26.8	23.1	34269	4	AAK68677	Aak68677 Human imm
44	26.8	23.1	34269	4	AAK85168	Aak85168 Human imm
C 45	26.6	22.9	2398	9	ACD19246	Ac19246 E. coli 0

#### ALIGNMENTS

RESULT 1

ABK46109

ID ABK46109 standard; CDNA; 116 BP.

XX AC ABK46109;

XX DT 05-JUN-2002 (first entry)

XX DE CDNA encoding colon tumour protein, SEQ ID NO 1660.

XX KW Human; colon tumour; vaccine; colon cancer; immunogenic; immunotherapy;

XX KW gene; ss.

XX OS Homo sapiens.

XX PN WO200212328-A2.

XX PD 14-FEB-2002.

XX PF 31-JUL-2001; 2001WO-US024218.

XX PR 03-AUG-2000; 2000US-0223283P.

XX PR 28-MAR-2001; 2001US-0279763P.

XX PR 29-JUN-2001; 2001US-0302051P.

XX (CORI-) CORIXA CORP.

XX King GE, Meagher MJ, Xu J, Secrist H;

XX WPI; 2002-241739/29.

XX PT New colon cancer polypeptides and polynucleotides, useful as vaccines, for

XX PT for diagnosing, preventing, and treating colon cancer, and as markers for

XX PT the progression of cancer.

XX PS Claim 1; SEQ ID NO 1660; 147bp; English.

XX PS The invention relates to polynucleotides encoding colon tumour proteins.

XX PS The polynucleotides and encoded polypeptides are useful in pharmaceutical

XX PS compositions, such as vaccines, for the diagnosis, prevention, and

XX PS treatment of colon cancer. Polynucleotide sequences may be used as

XX PS hybridisation probes or primers, and in the design and preparation of

XX PS ribozyme molecules for inhibiting expression of tumour polypeptides and

CC proteins in tumour cells. The compositions are useful for stimulating an  
CC immune response against cancer, particularly for the immunotherapy of  
CC colon cancer, and as markers for the progression of cancer. ABK4450-  
CC ABK46237 represent coding sequences of human colon tumour proteins of the  
CC invention. Note: With the exception of SEQ ID No 1 and 2, the sequence  
CC data for this patent did not form part of the printed specification but  
CC was supplied by the European Patent Office

XX  
XX  
SQ Sequence 116 BP; 34 A; 18 C; 35 G; 29 T; 0 U; 0 Other;

Query Match 100.0%; Score 116; DB 6; Length 116;  
Best Local Similarity 100.0%; Pred. NO. 1.3e-32;  
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGTAAAAATGACTCTTTCAGGAGTGTTCAGTAGGGTCAGATGACCACTGATTGGAA 60

DB 1 CAGGTAAAAATGACTCTTTCAGGAGTGTTCAGTAGGGTCAGATGACCACTGATTGGAA 60

QY 61 TACTTCGTAGCAGGAGCAAGTAGTCTGAGCCACTGTTCTATCGGTAGGGTGTC 116

DB 61 TACTTCGTAGCAGGAGCAAGTAGTCTGAGCCACTGTTCTATCGGTAGGGTGTC 116

## RESULT 2

AAA78132/c

ID AAA78132 standard; cDNA; 297 BP.

AC AAA78132;

DT 14-NOV-2000 (first entry)

DE cDNA encoding human colon tumour polypeptide, SEQ ID NO:419.

XX Human colon tumour polypeptide; tumour antigen; cancer; vaccine;

KW immunotherapy; diagnosis; progression; ss.

XX Homo sapiens.

XX WO200037643-A2.

XX 29-JUN-2000.

XX 23-DEC-1999; 99WO-US030909.

XX 23-DEC-1998; 98US-00221298.

XX 02-JUL-1999; 99US-00347496.

XX 22-SEP-1999; 99US-00401064.

XX 19-NOV-1999; 99US-00444242.

XX 02-DEC-1999; 99US-00454150.

XX (CORI-) CORIXA CORP.

XX Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stolk J;

PI Wang T, Yuqiu J;

XX WPI; 2000-442671/38.

XX New colon tumor polypeptides used to inhibit the development of cancer,

PT especially colon cancer, and for diagnosing and monitoring the

PT progression of the cancer.

XX Claim 1; Page 212; 229pp; English.

XX Sequences AAA7722-A78199 represent 478 cDNAs encoding proteins or

CC portions of proteins which are associated with human colon tumours. The

CC invention also specifically discloses 8 human colon tumour proteins

CC (AAB11897-B11904). The nucleic acids, the polypeptides they encode, and

CC antigen presenting cells (APCs), preferably dendritic cells, expressing

CC such polypeptides may be used in vaccines that target tumour cells,

CC especially colon tumour cells, thereby inhibiting the development of

CC cancer. T-cells specific for the polypeptide expressed by the APC are

CC used to remove tumour cells from biological samples, especially blood or

CC fractions thereof. The sample or the isolated T-cells specific for the

CC polypeptide can then be used to inhibit cancer development. CD4+ and/or  
CC CD8+ T-cells from a patient may be incubated with a polypeptide or  
CC nucleic acid of the invention, or an APC expressing such a polypeptide,  
CC to cause the proliferation of specific T-cells. The T-cells can be cloned  
CC and then administered back to the patient to inhibit cancer development.  
CC Nucleic acids encoding the polypeptides and antibodies against the  
CC polypeptides may be used to determine the expression level of a tumour  
CC protein of the invention, and therefore to determine whether cancer cells  
CC are present. Such diagnostic methods may also be used to monitor the  
CC progression of a cancer by repeating the processes at time intervals, and  
CC comparing the current result to previous results. The present sequence  
CC represents a cDNA encoding a human colon tumour polypeptide

XX  
SQ Sequence 297 BP; 84 A; 71 C; 41 G; 101 T; 0 U; 0 Other;

Query Match 94.3%; Score 109.4; DB 3; Length 297;

Best Local Similarity 99.1%; Pred. NO. 4.9e-30;

Matches 110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 AAAAAATGACTCTTTCAGGAGTGTTCAGTAGGGTCAGATGACCACTGATTGGAACTT 65

DB 136 AAAAAATGACTCTTTCAGGAGTGTTCAGTAGGGTCAGATGACCACTGATTGGAACTT 77

QY 66 CGTAAGCAGGAGCAAGTAGTCTGAGCCACTGTTCTATCGGTAGGGTGTC 116

DB 76 CGTAAGCAGGAGCAAGTAGTCTGAGCCACTGTTCTATCGGTAGGGTGTC 26

## RESULT 3

AAI28870/c

ID AAI28870 standard; cDNA; 297 BP.

AC AAI28870;

XX 12-OCT-2001 (first entry)

XX Colon tumour related determined cDNA sequence for clone 32020.

XX Human; immunotherapy; diagnosis; colon cancer; colon tumour; immunogenic;

KW Gene therapy; vaccine; colonic cancer; ss.

XX Homo sapiens.

XX WO200149716-A2.

XX 12-JUL-2001.

XX 29-DEC-2000; 2000WO-US035596.

XX 30-DEC-1999; 99US-00476296.

XX 10-JAN-2000; 2000US-00480321.

XX 15-FEB-2000; 2000US-00504629.

XX 06-MAR-2000; 2000US-00519444.

XX 19-MAY-2000; 2000US-00575251.

XX 29-JUN-2000; 2000US-00609448.

XX 28-AUG-2000; 2000US-00649811.

XX (CORI-) CORIXA CORP.

XX Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stolk JA;

PI King GE, Wang T, Jiang Y;

XX WPI; 2001-441847/47.

XX Colon tumor associated proteins and nucleic acids useful for the

PT prevention, diagnosis and treatment of colonic cancer.

XX Claim 2; Page 248; 472pp; English.

XX The present invention describes colon tumour associated proteins (I) and

CC the polynucleotides (II) that encode them. (I) have cytostatic activity.

CC (I) and (II) can be used in gene therapy and vaccine production. (I) and

CC (II) may be used in the prevention, diagnosis and treatment of diseases



```
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2002
/ LENGTH: 3579
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-2002

Query Match          94.3%; Score 109.4; DB 4; Length 3579;
Best Local Similarity 99.1%; Pred. No. 5e-31; 1; Indels 0; Gaps 0;
Matches 110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 AAAAAATGACTGTTTCAGGAGTGTTCAGTAGGTCAGATGACCAAGTATGGGAATACCTT 65
DB 3414 AAAACATGACTGTTTCAGGAGTGTTCAGTAGGTCAGATGACCAAGTATGGGAATACCTT 3355

QY 66 CGTAAGCAGGAGCAAGTAAGATCTGAGCCACTGTTCTATCGGTAGGGTCTC 116
DB 3354 CGTAAGCAGGAGCAAGTAAGATCTGAGCCACTGTTCTATCGGTAGGGTCTC 3304

RESULT 3
US-09-949-016-13744/c
/ Sequence 13744, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 13744
/ LENGTH: 21800
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(21800)
/ OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13744

Query Match          94.3%; Score 109.4; DB 4; Length 21800;
Best Local Similarity 99.1%; Pred. No. 1e-30; 1; Indels 0; Gaps 0;
Matches 110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 AAAAAATGACTGTTTCAGGAGTGTTCAGTAGGTCAGATGACCAAGTATGGGAATACCTT 65
DB 19634 AAAACATGACTGTTTCAGGAGTGTTCAGTAGGTCAGATGACCAAGTATGGGAATACCTT 19575

QY 66 CGTAAGCAGGAGCAAGTAAGATCTGAGCCACTGTTCTATCGGTAGGGTCTC 116
DB 19574 CGTAAGCAGGAGCAAGTAAGATCTGAGCCACTGTTCTATCGGTAGGGTCTC 19524

RESULT 4
US-09-248-796A-11178
/ Sequence 11178, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Keith Weinstock et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ CURRENT FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 11178
/ LENGTH: 198
/ TYPE: DNA
/ ORGANISM: Candida albicans
US-09-248-796A-11178

Query Match          24.1%; Score 28; DB 4; Length 198;
Best Local Similarity 55.0%; Pred. No. 0.84;
Matches 55; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 1 CAGGTAAAAATGACTGTTTCAGGAGTGTTCAGTAGGTCAGATGACCAAGTATGGGA 60
DB 62 CAAAAAAGAAAAATGATGAAGCGGTATCAAAAGTACAAGCCAAAAATCCCTCATTTGGA 121

QY 61 TACTTCGTAAGCAGGAGCAAGTAAGATCTGAGCCACTGTT 100
DB 122 GAATACATGAGCATAAACAATCACACATTTGTGCAAAATGTT 161

RESULT 5
US-09-949-016-11835
/ Sequence 11835, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 11835
/ LENGTH: 100550
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-11835

Query Match          23.6%; Score 27.4; DB 4; Length 100550;
Best Local Similarity 65.6%; Pred. No. 16;
Matches 40; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 15 CTGTTTCAGGAGTGTTCAGTAGGTCAGATGACCAAGTATGGGAATACCTTCGTAAGCAG 74
DB 3734 CTTTGCAACAGTGACAGAGAAGGGTCAGTCGACCAATGGTCTGGAACTTCTGTATAGGCAG 3793

QY 75 G 75
DB 3794 G 3794

RESULT 6
US-09-949-016-16207
/ Sequence 16207, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
```